

Article

The COVID-19 Mortality Rate in Latin America: A Cross-Country Analysis

Fernando José Monteiro de Araújo ^{1,*}, Renata Rojas Guerra ^{1,2} and Fernando Arturo Peña-Ramírez ²¹ Graduate Program of Statistics, Universidade Federal do Rio Grande do Sul, Porto Alegre 90010-150, Brazil² Department of Statistics, Universidade Federal de Santa Maria, Santa Maria 97105-900, Brazil; renata.r.guerra@ufsm.br (R.R.G.); fernando.p.ramirez@ufsm.br (F.A.P.-R.)

* Correspondence: fernando.monteiro@ufrgs.br

Abstract: Latin America was one of the hotspots of COVID-19 during the pandemic. Therefore, understanding the COVID-19 mortality rate in Latin America is crucial, as it can help identify at-risk populations and evaluate the quality of healthcare. In an effort to find a more flexible and suitable model, this work formulates a new quantile regression model based on the unit ratio-Weibull (URW) distribution, aiming to identify the factors that explain the COVID-19 mortality rate in Latin America. We define a systematic structure for the two parameters of the distribution: one represents a quantile of the distribution, while the other is a shape parameter. Additionally, some mathematical properties of the new regression model are presented. Point and interval estimates of maximum likelihood in finite samples are evaluated through Monte Carlo simulations. Diagnostic analysis and model selection are also discussed. Finally, an empirical application is presented to understand and quantify the effects of economic, social, demographic, public health, and climatic variables on the COVID-19 mortality rate quantiles in Latin America. The utility of the proposed model is illustrated by comparing it with other widely explored quantile models in the literature, such as Kumaraswamy and unit Weibull regressions.

Keywords: epidemiology; quantiles; unit extended Weibull families; unit regression**MSC:** 62J20; 62E10

Citation: de Araújo, F.J.M.; Guerra, R.R.; Peña-Ramírez, F.A. The COVID-19 Mortality Rate in Latin America: A Cross-Country Analysis. *Mathematics* **2024**, *12*, 3934. <https://doi.org/10.3390/math12243934>

Academic Editors: Domma Filippo and Francesca Condino

Received: 21 November 2024

Revised: 8 December 2024

Accepted: 10 December 2024

Published: 13 December 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is the etiological agent of coronavirus disease 2019 (COVID-19). On 11 March 2020, the WHO declared COVID-19 a global pandemic [1]. Approximately a year and two months after the WHO declared the COVID-19 epidemic a global public health emergency, on 23 March 2021, Brazil reported 3158 deaths in just 24 h. This figure was the highest number of deaths reported worldwide since the beginning of the pandemic. That same day, Argentina appeared in the top 10 countries with the highest number of daily infections (9405), and Colombia occupied the 11th position in total deaths (62,274) since the beginning of the emergency. Peru was also cataloged as the country in the Latin American region with more deaths per 1000 inhabitants (0.516).

In May 2020, the World Health Organization (WHO) declared that South America had become the new epicenter of the COVID-19 pandemic, with countries like Brazil, Argentina, and Peru reporting some of the highest per capita mortality rates worldwide [2]. As the pandemic progressed, Latin America became again an epicenter in September 2020. Although the United States continued to lead globally in total cases and deaths, Brazil ranked second, followed by other heavily affected countries in the region, such as Peru, Chile, Mexico, Colombia, Ecuador, Argentina, and Bolivia [3]. Between March and June 2021, the epicenter shifted back to Latin America, driven by the spread of the Delta variant,

which quickly became the dominant strain globally [4]. The prevalence of comorbidities such as diarrhea and diabetes leaves the region in a complex and delicate clinical and epidemiological environment. The situation worsens with the coexistence of other epidemics, such as dengue and yellow fever, in addition to the long-term consequences of chikungunya and Zika [5].

In the context of epidemic modeling, predicting infections and mortality rates at a regional or national level is essential. [6] discusses the use of compartmental models such as susceptible-infected-recovered (SIR) and its variants, including susceptible-exposed-infected-recovered (SEIR) and susceptible-infected-susceptible (SIS), among others. The SIS model, initially proposed by [7], is a mathematical model developed to describe the dynamics of epidemics, such as infection rate and population immunity, and quantifies how an infectious disease spreads over time. These models remain widely used, as demonstrated by [8], who applied the SIR model to predict the number of cases of COVID-19 in Malaysia during different pandemic phases.

Considering probabilistic modeling, several authors, such as [9,10], have investigated how climatic and cultural factors influence the death rate from COVID-19. In such cases, classic linear regression models have been fitted to explore these associations, accounting for factors like temperature, humidity, and cultural dimensions. It is common to apply this methodology to explain mortality rates based on other variables. Additionally, recent works by [11,12] have utilized regression models to analyze the impacts of factors such as government effectiveness, testing rates, and public health measures on COVID-19 mortality. However, this methodology brings some limitations, particularly the assumption of normal distribution in the response variable. This assumption does not capture the character of the response variable since the rates are bounded random variables and, most of the time, asymmetric. The normal distribution cannot represent this characteristic, and we note that the specialized literature has paid little attention to this fact.

A modeling alternative is to assume a distribution in the exponential family for the response variable. These are the well-known generalized linear models (GLMs) pioneered by [13]. However, this assumption remains restrictive. Another class of more general and flexible models assumes that the response distribution is deliberately left general with no explicit distribution specified, and its parameters vary as a function of explanatory variables. They are the Generalized Additive Models for Location, Scale, and Shape (GAMLSS) framework [14]. Recently, applications and proposals for models based on the GAMLSS approach have gained prominence. For example, in the study by [15], a variety of models were used, including beta, simplex, unit gamma (UG), and unit Lindley (UL) regressions, to identify covariates associated with the proportion of votes in municipal elections.

In this paper, we directly utilize the GAMLSS framework to formulate a new regression model based on the Weibull distribution, aimed at explaining the COVID-19 mortality rate in Latin America. We define a structure of systematic components on the two parameters of the distribution: one of which represents a quantile of the distribution and the other its dispersion. Modeling on the median is preferable over the mean when the variable of interest is not symmetric (i.e., skewed), especially in the presence of outliers [16], in addition to being a more robust measure of central tendency [17].

The main contribution of this paper is to propose a new regression model that facilitates understanding and quantifying the impact of economic variables, social and demographic indicators, and public health measures on the quantiles of the COVID-19 mortality rate. Unlike commonly used models for these purposes, this model accommodates the typical asymmetry and bounded nature of mortality rate data. We hope that our approach can serve as a valuable tool for policymakers in decision-making. We focus on the initial mortality rates of the pandemic, as other variables, such as government responses through public policies promoting mask-wearing, widespread testing, and social isolation, began to influence these rates as the pandemic advanced. Such initiatives underscore the importance of government interventions in public health to mitigate the pandemic's impacts.

The model can be effectively applied to analyze mortality rates of diseases with low fatality rates, such as measles, or other epidemic diseases, such as dengue and yellow fever. This model can be extended to analyze various economic and engineering applications. In economics, it can model data like the Gini index and poverty rates, which are proportions between 0 and 1 and often exhibit positive skewness. In engineering, the model can be used to analyze failure rates of systems or efficiency metrics, which may also be expressed as percentages (ranging from 0 to 100%) and are typically skewed. These applications highlight the importance of using more flexible models to interpret complex, asymmetrical data.

The remainder of the paper is outlined as follows. Section 2 introduces the new regression model, the unit ratio-Weibull distribution for the COVID-19 mortality rates in Latin American countries. Further estimation and goodness-of-fit aspects are also presented. In Section 3, a Monte Carlo simulation study is conducted to evaluate the performance of the maximum likelihood estimators of the proposed regression model. Section 4 describes the data preparation and carries out the regression analysis by comparing the novel model with other quantile regressions in the unit interval. The concluding remarks are addressed in Section 5.

2. The Unit Ratio-Weibull Regression

This section introduces a new quantile regression for modeling double-bounded epidemiological data. By focusing on COVID-19 applications, our approach arises as an alternative to analyze the impact of demographic and epidemiological indicators on the mortality rate of this disease. The proposed regression is based on the unit ratio-Weibull (URW) distribution, which belongs to the unit ratio-extended Weibull family and was pioneered by [18].

A random variable Y has a URW distribution, denoted by $Y \sim URW(\sigma, \mu)$, if its cumulative distribution function (cdf) and probability density function (pdf) are

$$F(y|\sigma, \mu) = 1 - (1 - \tau)^{y^\sigma (1-\mu)^\sigma / [\mu^\sigma (1-y)^\sigma]}, \quad y \in (0, 1) \tag{1}$$

and

$$f(y|\sigma, \mu) = \frac{\sigma y^{\sigma-1} (1 - \mu)^\sigma}{\mu^\sigma (1 - y)^{\sigma+1}} \log \left[(1 - \tau)^{-1} \right] (1 - \tau)^{y^\sigma (1-\mu)^\sigma / [\mu^\sigma (1-y)^\sigma]}, \tag{2}$$

respectively, where $\sigma > 0$ is a shape parameter, and $\mu \in (0, 1)$ is the τ th quantile of the distribution. For $\sigma = 2$, the unit ratio-Rayleigh (URR) distribution yields a special case. The corresponding quantile function (qf) is

$$Q(u|\sigma, \mu) = \frac{\left[\frac{\mu^\sigma \log(1-u)}{(1-\mu)^\sigma \log(1-\tau)} \right]^{1/\sigma}}{1 + \left[\frac{\mu^\sigma \log(1-u)}{(1-\mu)^\sigma \log(1-\tau)} \right]^{1/\sigma}}. \tag{3}$$

Since $Q(\tau|\sigma, \mu) = \mu$, it follows that μ is a location parameter corresponding to the URW τ th quantile.

The flexibility and advantages of the URW distribution and its special case are illustrated by numerical experiments in real and simulated data sets [18]. However, the URW does not accommodate explanatory variables in the modeling. The current paper overcomes this limitation by introducing the URW regression in which the quantile and the shape parameters can be related to a linear predictor. The useful parameterization of the URW allows us to formulate a quantile regression model that consists of two components, namely:

- (i) a random component in which Y_1, \dots, Y_n is a sample of n independent random variables, where each $Y_t, t = 1, \dots, n$, follows a URW distribution with quantile μ_t and shape parameter σ_t , that is, $Y_t \sim URW(\sigma_t, \mu_t)$;

(ii) systematic components through the linear predictors

$$\eta_{1t} = g_1(\mu_t) = \mathbf{x}_t^\top \boldsymbol{\beta}, \tag{4}$$

and

$$\eta_{2t} = g_2(\sigma_t) = \mathbf{z}_t^\top \boldsymbol{\gamma}, \tag{5}$$

where $\mathbf{x}_t = (1, x_{t2}, \dots, x_{tk})^\top$ and $\mathbf{z}_t = (1, z_{t2}, \dots, z_{tl})^\top$ are $k \times 1$ and $l \times 1$ vectors that contain observations of k and l known covariates ($k + l < n$), respectively. The vectors of unknown regression parameters are $\boldsymbol{\beta} = (\beta_1, \dots, \beta_k)^\top \in \mathbb{R}^k$ and $\boldsymbol{\gamma} = (\gamma_1, \dots, \gamma_l)^\top \in \mathbb{R}^l$. Finally, $g_1 : (0, 1) \rightarrow \mathbb{R}$ and $g_2 : \mathbb{R}^+ \rightarrow \mathbb{R}$ are strictly monotonic and twice differentiable link functions that differ on the mapping required. Ref. [13] provides an overview of some classical link functions under the generalized linear models approach.

2.1. Parameter Estimation

Several approaches to parameter estimation are explored in the literature. However, given its desirable and known asymptotic properties, the maximum likelihood method is the most widely used. In this section, we determine the maximum likelihood estimators (MLEs) of the parameters of the URW regression. Let $\boldsymbol{\theta}^\top = (\boldsymbol{\beta}^\top, \boldsymbol{\gamma}^\top)$ be the URW regression parameter vector, and $\mathbf{y}^\top = (y_1, \dots, y_n)$ the corresponding sample of n independent observations. The log-likelihood function for this sample is

$$\begin{aligned} \ell_t(\sigma_t, \mu_t) &= \sum_{t=1}^n \log\left(\frac{\sigma_t}{y_t}\right) + \sum_{t=1}^n \log\left[\frac{\log(1 - \tau)}{y_t - 1}\right] + \sum_{t=1}^n \log\left[\frac{y_t(1 - \mu_t)}{\mu_t(1 - y_t)}\right]^{\sigma_t} \\ &+ \log(1 - \tau) \sum_{t=1}^n \left[\frac{y_t(1 - \mu_t)}{\mu_t(1 - y_t)}\right]^{\sigma_t}, \end{aligned} \tag{6}$$

where $\mu_t = g_1^{-1}(\eta_{1t})$, and $\sigma_t = g_2^{-1}(\eta_{2t})$, with η_{1t} and η_{2t} given in (4) and (5), respectively. The score function is obtained by $\mathbf{U} = (\mathbf{U}_\beta(\boldsymbol{\beta}, \boldsymbol{\gamma})^\top, \mathbf{U}_\gamma(\boldsymbol{\beta}, \boldsymbol{\gamma})^\top)^\top$, where

$$\mathbf{U}_\beta(\boldsymbol{\theta}) = \partial \ell(\boldsymbol{\theta}) / \partial \boldsymbol{\beta}^\top = \mathbf{X}^\top \mathbf{T} \mathbf{w},$$

with \mathbf{X} is a $n \times k$ matrix whose t th row is \mathbf{x}_t^\top , $\mathbf{T} = \text{diag}\{1/g'_1(\mu_1), \dots, 1/g'_1(\mu_n)\}$, $\mathbf{w} = (w_1, \dots, w_n)$ wherein

$$w_t = \frac{\sigma_t}{\mu_t(\mu_t - 1)} [1 + \log(1 - \tau)k_t^{\sigma_t}],$$

$k_t = y_t(\mu_t - 1) / \mu_t(y_t - 1)$, and $g'_1(\mu_t)$ is the differentiating of $g_1(\mu_t)$ with respect to μ_t , and

$$\mathbf{U}_\gamma(\boldsymbol{\theta}) = \partial \ell(\boldsymbol{\theta}) / \partial \boldsymbol{\gamma}^\top = \mathbf{Z}^\top \mathbf{S} \mathbf{u},$$

wherein \mathbf{Z} is a $n \times l$ matrix whose t th row is \mathbf{z}_t^\top , $\mathbf{S} = \text{diag}\{1/g'_2(\sigma_1), \dots, 1/g'_2(\sigma_n)\}$, $\mathbf{u} = (u_1, \dots, u_n)$ with

$$u_t = \frac{1}{\sigma_t} + \log(k_t) [1 + \log(1 - \tau)k_t^{\sigma_t}],$$

and $g'_2(\sigma_t)$ is the differentiating of $g_2(\sigma_t)$ with respect to σ_t .

The Hessian matrix is given by

$$K(\theta) = \frac{\partial^2 \ell(\theta)}{\partial \theta \partial \theta^\top} = \begin{pmatrix} \frac{\partial^2 \ell(\theta)}{\partial \beta^2} & \frac{\partial^2 \ell(\theta)}{\partial \beta \partial \gamma^\top} \\ \frac{\partial^2 \ell(\theta)}{\partial \gamma \partial \beta^\top} & \frac{\partial^2 \ell(\theta)}{\partial \gamma^2} \end{pmatrix},$$

where

$$\frac{\partial^2 \ell(\theta)}{\partial \beta^2} = \mathbf{X}^\top \mathbf{M} \mathbf{X},$$

$$\frac{\partial^2 \ell(\theta)}{\partial \gamma^2} = \mathbf{Z}^\top \mathbf{B} \mathbf{Z},$$

$$\frac{\partial^2 \ell(\theta)}{\partial \beta \partial \gamma^\top} = \frac{\partial^2 \ell(\theta)}{\partial \gamma \partial \beta^\top} = \mathbf{Z}^\top \mathbf{D} \mathbf{X},$$

and $\mathbf{M} = \text{diag}\{m_1, \dots, m_n\}$, $\mathbf{B} = \text{diag}\{b_1, \dots, b_n\}$, $\mathbf{D} = \text{diag}\{d_1, \dots, d_n\}$, wherein

$$m_t = \frac{\sigma_t}{\mu_t^2(\mu_t - 1)^2} [1 - 2\mu_t + k_t^{\sigma_t} \log(1 - \tau)(1 - 2\mu_t + \sigma_t)] \left[\frac{1}{g'_1(\mu_t)} \right]^2,$$

$$b_t = \left[k_t^{\sigma_t} \log^2(k_t) \log(1 - \tau) - \frac{1}{\sigma_t^2} \right] \left[\frac{1}{g'_2(\sigma_t)} \right]^2$$

and

$$d_t = \frac{1}{\mu_t(\mu_t - 1)} \{1 + k_t^{\sigma_t} \log(1 - \tau)[1 + \sigma_t \log(k_t)]\} \left[\frac{1}{g'_1(\mu_t)} \right] \left[\frac{1}{g'_2(\sigma_t)} \right].$$

Under regularity conditions, the asymptotic normality property of EMVs ensures that when the sample size is large,

$$\begin{pmatrix} \hat{\beta} \\ \hat{\gamma} \end{pmatrix} \sim N_{k+l} \left(\begin{pmatrix} \beta \\ \gamma \end{pmatrix}, [-K(\theta)]^{-1} \right),$$

approximately. Moreover, $[-K(\theta)]^{-1}$ is the asymptotic variance–covariance matrix of $\hat{\theta}$, and $-K(\theta)$ is the observed information matrix.

To determine the MLEs of θ , denoted as $\hat{\theta} = (\hat{\beta}, \hat{\gamma})^\top$, is necessary to maximize (6) by setting the score vector components at zero and solving the system of equations simultaneously. However, that is a non-linear system, and numerical methods must be used. Since the proposed model resembles the Generalized Additive Models for Location, Scale, and Shape (GAMLSS) approach [14,19], we implement the URW regression as a gamlss class object in R programming language, which is available in the gamlss package [20,21] and uses the RS algorithm for maximizing the log-likelihood given in (6). The computational codes for the URW model and the simulation and application studies can be downloaded from <https://github.com/Fernando-code8/URW-Unit-Ratio-Weibull-Regression> (accessed on 18 October 2024). The cdf (1), pdf (2), and qf (3) can be computed using the dURW, pURW, and qURW functions, respectively. Samples of the URW model can be generated using the rURW function.

2.2. Diagnostic Measures and Model Selection

Diagnostic measures are customarily adopted to check if a fitted regression model adequately represents the data dynamics. To that end, we perform residual analysis using the quantile residuals introduced by [22]. For the URW regression, such a residual is given by $r_t = \Phi^{-1}[F(y_t|\hat{\sigma}_t, \hat{\mu}_t)]$, where $F(\cdot|\hat{\sigma}_t, \hat{\mu}_t)$ is obtained from the URW cdf (1) evaluated at $\hat{\mu}_t = g_1^{-1}(x_t^\top \hat{\beta})$ and $\hat{\sigma}_t = g_2^{-1}(z_t^\top \hat{\gamma})$. In the literature, several authors have been considering the quantile residuals since they are standard normally distributed when the model is correctly specified [22]. See [17,23,24] for instance.

We also consider the worm plot of the residuals to verify whether the assumed distribution fits properly for the dependent variable [25]. We expect that $100(1 - \alpha)\%$ of the points to be inside the two elliptic curves in the middle of the figure. A large proportion of points outside this region and the occurrence of any specific shape in the points indicate that the fitted model is inadequate.

The generalized coefficient of determination (R_G^2) is considered to measure the predictive capacity of URW-fitted regressions. Defined by [26], the R_G^2 is given by $R_G^2 = 1 - \exp\{-2/n[\ell(\hat{\theta}) - \ell(\hat{\theta}_0)]\}$, where $\ell(\hat{\theta})$ is the log-likelihood of the fitted model, and $\ell(\hat{\theta}_0)$ is the log-likelihood of the model without covariates, i.e., the null model. The higher the R_G^2 , the better the fitted model to explain the variability of the response variable.

Finally, Akaike information criteria (AIC) are suggested for model selection. The AIC is widely used to select the more suitable model among a class of candidate models and is defined by [27] as $AIC = 2[m - \ell(\hat{\theta})]$, where $m = k + l$ is the number of estimated parameters. The better model is the one with a smaller AIC.

The model performance was assessed using leave-one-out cross-validation (LOOCV); see [28] for details. This methodology involves sequentially partitioning the dataset into two parts. Consider a dataset $\{(y_1, x_1), \dots, (y_n, x_n)\}$ consisting of responses Y and their associated vectors of k covariates, $x_i = (x_{i1}, \dots, x_{ik})$, where $i = 1, \dots, n$. Let \hat{y}_i^* be the estimate of the value y_i , obtained by excluding the i -th observation from the fit. Specifically, this involves fitting a regression model to the dataset $\{(y_1, x_1), \dots, (y_{i-1}, x_{i-1}), (y_{i+1}, x_{i+1}), \dots, (y_n, x_n)\}$ and then substituting x_i into the fitted regression structure to compute the estimate of y_i , denoted by \hat{y}_i^* . This procedure is repeated for all observations in the dataset to obtain y_1^*, \dots, y_n^* .

To compare the predicted values of the fitted models, we use the mean absolute error (MAE), defined as $MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i^*|$. MAE quantifies the absolute difference between observed and predicted values. Therefore, a lower MAE indicates better model performance.

3. Numerical Evidence

In what follows, we report Monte Carlo experiments to explore the performance of the maximum likelihood method and the assumptions on the empirical distribution of r_t for the proposed methodology. We generate 10,000 replications of a URW regression with the systematic components given by (4) and (5) for the quantile and shape parameters, respectively. The logit link function is used for μ , $\text{logit}(\mu) = \log[\mu/(1 - \mu)]$, and the log link function is used for σ . The sample sizes are set at $n \in \{10, 15, 30, 70, 150, 300\}$ and the values of the covariates are generated from the standard uniform distribution. For the parameter values, two different scenarios are considered, namely:

- Scenario 1: $\beta_1 = -3.75$, $\beta_2 = 0.25$, $\gamma_1 = 1.5$, and $\gamma_2 = 1.5$;
- Scenario 2: $\beta_1 = -5$, $\beta_2 = 1.75$, $\gamma_1 = 2$, and $\gamma_2 = 1.25$.

For brevity, we will present the results with $\tau = 0.5$, which represents the median of the response variable. The numerical evidence for other quantiles was quite similar and is reported in the Appendix A.

The measurements calculated to evaluate the point estimators are the percentage relative bias (RB%), mean square error (MSE), coefficient of skewness (CS), and kurtosis (K). We also compute the coverage rate (CR%) for the interval estimators with a nominal level

at 0.95. The results are summarized in Table 1. Notice the point estimators become more accurate and precise as the sample size increases; for example, when $n = 300$, the MSE is less than 0.15 for all parameter estimates and scenarios. The CS and K coefficients are in line with the expected since they get closer to 0 and 3, respectively, as n increases. When analyzing the CR%, we verify values close to the nominal level. The largest difference, of only 0.1248, occurs for $\hat{\beta}_1$ with the smallest sample size of $n = 10$ in Scenario 1. This still represents a minor discrepancy, particularly given the small sample size.

We evaluate the mean, variance, CS, and K measures to study the empirical distribution of the quantile residuals. We expect those statistics to be close to 0, 1, 0, and 3, respectively, since the quantile residual distribution is expected to be approximately standard normal. In this sense, we also compute the null rejection rates for the Shapiro–Wilk normality test at the 10%, 5%, and 1% significance levels, which are referred to as NRR10%, NRR5%, and NRR1%, respectively. Table 2 reports the results for the residual simulations. The calculated statistics show that the distribution of r_i is approximated by its reference distribution. The normality test corroborates this result since the null rejection rates are close to the test nominal level for all scenarios.

Table 1. Results of simulation of URW regression with $\tau = 0.5$.

Measures	n	Scenario 1				Scenario 2			
		$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\gamma}_1$	$\hat{\gamma}_2$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\gamma}_1$	$\hat{\gamma}_2$
RB%	10	0.4321	−9.4939	66.3225	−13.6220	0.1536	−2.2519	55.3137	−17.9652
	15	0.0532	−5.1174	27.2422	17.6428	0.6361	1.9085	28.6735	7.1978
	30	0.0342	−1.0949	13.7665	4.5770	0.0308	−0.0859	13.5165	2.1807
	70	0.0036	−0.1549	3.9841	5.0650	0.0000	−0.0153	3.9081	5.8399
	150	−0.0048	−0.3322	1.7043	2.5107	−0.0033	−0.0404	1.7326	2.7387
	300	−0.0055	−0.5761	1.0077	0.8726	−0.0044	−0.0708	0.9927	0.8967
MSE	10	0.0858	0.4924	9.8697	22.0775	0.0644	0.3705	13.0225	28.6075
	15	0.0691	0.2481	1.2178	5.1895	0.1011	0.1814	2.2252	5.8772
	30	0.0429	0.1130	0.4112	1.3282	0.0306	0.0827	0.6725	1.8748
	70	0.0130	0.0326	0.1368	0.5958	0.0097	0.0246	0.2164	0.8274
	150	0.0055	0.0181	0.0391	0.1932	0.0041	0.0134	0.0637	0.2644
	300	0.0027	0.0072	0.0232	0.1047	0.0020	0.0054	0.0370	0.1441
CS	10	−0.4821	0.0163	1.3254	−0.0668	−0.4807	0.0217	1.3911	−0.0769
	15	−0.3653	−0.0450	1.6770	0.1168	−0.4961	0.2488	1.8602	−0.2006
	30	−0.2080	0.0287	1.4250	−0.0677	−0.2165	0.0293	1.2909	−0.1013
	70	−0.1485	0.0161	0.6854	0.0520	−0.1490	0.0167	0.6490	0.0402
	150	−0.1197	0.0100	0.4637	0.0724	−0.1187	0.0134	0.4418	0.0514
	300	−0.1205	0.0032	0.3357	0.0099	−0.1185	0.0037	0.3177	−0.0048
K	10	3.7022	3.4739	6.1684	4.0920	3.6177	3.4202	6.9478	4.2937
	15	3.4778	3.3128	7.2405	4.4532	3.6312	3.4633	8.6702	4.8173
	30	3.1766	3.2517	6.8859	4.0104	3.1935	3.2596	6.0387	3.9389
	70	3.1527	3.1266	3.9511	3.4472	3.1691	3.1243	3.8532	3.4277
	150	3.0272	2.9784	3.3380	3.2861	3.0159	2.9739	3.2966	3.2884
	300	3.0321	3.0402	3.1597	3.0193	3.0371	3.0379	3.1356	3.0045
CR%	10	0.8371	0.8252	0.9327	0.9308	0.8420	0.8303	0.9334	0.9324
	15	0.8709	0.8659	0.9365	0.9261	0.8661	0.8647	0.9404	0.9378
	30	0.9101	0.9128	0.9433	0.9374	0.9077	0.9131	0.9443	0.9375
	70	0.9360	0.9404	0.9487	0.9475	0.9378	0.9396	0.9474	0.9467
	150	0.9446	0.9408	0.9537	0.9505	0.9443	0.9409	0.9531	0.9497
	300	0.9479	0.9488	0.9476	0.9467	0.9474	0.9471	0.9482	0.9479

Table 2. Simulation results for the quantile residuals of the URW regression with $\tau = 0.5$.

Scenario	<i>n</i>	Mean	Variance	CS	K	NRR10%	NRR5%	NRR1%
1	10	0.0086	1.1376	−0.0801	2.2805	0.0747	0.0316	0.0059
	15	0.0022	1.0920	−0.0668	2.3691	0.0877	0.0379	0.0036
	30	−0.0011	1.0464	−0.0298	2.6328	0.0874	0.0391	0.0050
	70	−0.0005	1.0203	−0.0146	2.8228	0.0902	0.0436	0.0073
	150	−0.0003	1.0097	−0.0090	2.9047	0.0848	0.0405	0.0075
	300	0.0004	1.0050	−0.0054	2.9518	0.0909	0.0420	0.0077
2	10	0.0068	1.1398	−0.0871	2.2747	0.0745	0.0296	0.0047
	15	0.0032	1.0946	−0.0764	2.3614	0.0816	0.0340	0.0036
	30	−0.0015	1.0471	−0.0304	2.6306	0.0869	0.0389	0.0051
	70	−0.0008	1.0206	−0.0148	2.8220	0.0900	0.0428	0.0075
	150	−0.0003	1.0099	−0.0091	2.9042	0.0842	0.0410	0.0072
	300	0.0003	1.0051	−0.0054	2.9518	0.0904	0.0418	0.0079

4. An Analysis of COVID-19 Mortality Rate

The current section outlines the data preparation employed to investigate the COVID-19 mortality rate. We consider data from 19 Latin American countries: Argentina, Belize, Bolivia, Brazil, Chile, Colombia, Costa Rica, Ecuador, El Salvador, Guatemala, Honduras, Mexico, Nicaragua, Panama, Paraguay, Peru, Suriname, Uruguay, and Venezuela. For comparative purposes, the response variable is defined as the initial mortality rate (MR) per thousand people 90 days after the 20th detected case. We fit the URW regression concurrently with the Kumaraswamy [29] (KW) and unit Weibull [17] (UW) quantile regressions. Those are well-known unit regressions that may be alternatives when the interest lies in modeling the impact of explanatory variables in a quantile of the mortality rate.

We analyze the impact of pre-existing country characteristics such as social, demographic, and health indicators in the MR. To this aim, we consider the data from the most recent year available, which were collected from [30]. The selected explanatory variables are presented in Table 3. The rest of the section presents a descriptive summary and correlation analysis for the considered variables and the regression models’ results. This information can be helpful to reveal epidemiological differences across Latin American countries, identify the covariates with larger influence in the response variable, and understand their impact on the initial COVID-19 mortality rate. Finally, the results may contribute to shaping a direction for predicting regional or national infections and mortality in future research.

Table 3. Definition of the variables.

Variable	Description
CHE	Current health expenditure: Level of current health expenditure expressed as a percentage of the gross domestic product in the year. It includes healthcare goods and services consumed during each year 2017. This indicator does not include capital health expenditures such as buildings, machinery, and stocks of vaccines for emergencies or outbreaks.
DGGHE	Domestic general government health expenditure: Public expenditure on health from domestic sources as a share of the economy as measured by the percentage of the gross domestic product in 2017.
DP	Diabetes prevalence: percentage of people ages 20–79 who have type 1 or type 2 diabetes in the year 2019.
GDP	Gross domestic product per capita: gross domestic product divided by midyear population in 2019. The exceptions are Cuba and Venezuela, which the more recent information was for 2018 and 2014, respectively.
HDI	Human development index in 2018.
UP	Urban population: percentage of the total population of people living in urban areas in 2019.

4.1. Descriptive Summary and Correlation Analysis

The COVID-19 mortality rate had an asymmetrical and dispersed behavior among the Latin American countries; see its skewness and percentage coefficient of variation (CV%) in Table 4. From Figure 1, notice the MR is mostly concentrated on the left tail, indicating a right-skewed and decreasing shape on the histogram. Thus, the MR is heterogeneous among the Latin American countries, which is a typical feature for these kind of data. For example, [12] reported that the COVID-19 mortality rate varies greatly and has a right-skewed distribution across countries, and [3] verified that this disease manifests differently among the various regions of Latin America.

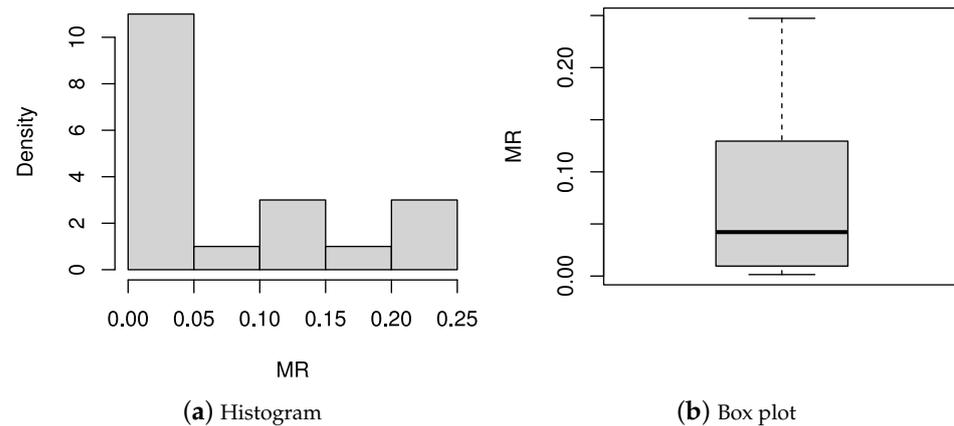


Figure 1. Histogram and box plot of the MR.

Table 4 also shows a descriptive summary for the explanatory variables. The highest variability is observed for the GDP, which has CV% around 240 and also presents the highest positive value of skewness. [31] highlights, through the Gini Index, how inequality among Latin American countries in relation to GDP has persisted over time. The HDI presents the lower CV% and most of the countries are classified with high, between 0.7–0.799 [32], or very high, between 0.8–1 [32]. Variables CHE, DGGHE, HDI, and UP have negative skewness. The variables MR, UP, and HDI, given their negative kurtosis, have light-tailed distributions. The variables with the heaviest-tailed distributions are PD and GDP.

Table 4. Descriptive summary for the response variable and covariates.

Variable	Mean	Median	Skewness	Kurtosis	Min.	Max.	CV (%)
CHE	7.0112	7.2276	−1.1365	1.5606	1.1812	9.4675	28.0252
DGGHE	4.0354	4.3616	−0.4765	0.4420	0.1883	6.6089	37.1126
DP	9.0789	8.6000	1.1076	0.6581	5.5000	17.1000	32.1934
GDP	3.2553	0.0226	2.3576	4.6840	0.0039	29.7482	239.1504
HDI	0.7405	0.7580	−0.2021	−0.9595	0.6230	0.8470	8.4062
MR	0.0793	0.0422	0.7314	−1.0656	0.0014	0.2473	109.1801
UP	72.9568	72.7460	−0.1786	−1.1724	45.8660	95.4260	19.4462

Table 5 presents the correlation matrix for the studied variables. The correlations presented are those corresponding to the Spearman method. We can observe negative correlations between the response variable (MR) and the GDP and DP variables. On the other hand, the highest positive correlation is that associated with the HB variable. Figure 2 displays scatterplots of the MR versus other covariates and provides a visual inspection of the correlation measure.

Table 5. Spearman’s correlation coefficient between all variables.

Variables	DGGHE	DP	GDP	HDI	MR	UP
CHE	0.6895	−0.1676	0.8939	0.3549	0.0947	0.2930
DGGHE		−0.2255	0.7924	0.4207	−0.1351	0.3842
DP			0.4178	−0.2241	−0.2159	−0.3668
GDP				0.2143	0.3130	0.2143
HDI					0.1880	0.7668
MR						0.0211

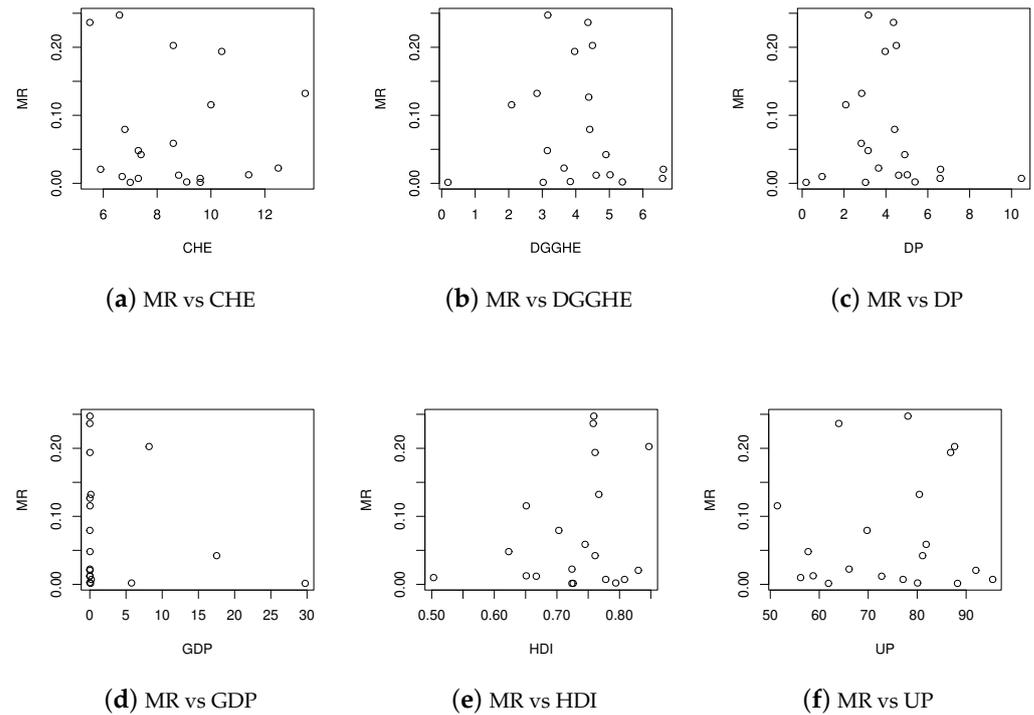


Figure 2. Dispersion plots of MR as a function of CHE, DGGHE, DP, GDP, HDI and UP.

4.2. Regression Results and Discussion

The COVID-19 mortality rate analysis is performed by taking the MR as the response variable in the proposed regression. Two competitor unit models are considered and compared to the URW regression. The Kumaraswamy and UW quantile regressions have their random components given by the pdfs

$$f(y|\sigma, \mu) = \frac{\log(1 - \tau)}{\sigma \log(1 - \mu^{1/\sigma})} y^{1/\sigma - 1} (1 - y^{1/\sigma})^{\log(1 - \tau) / \log(1 - \mu^{1/\sigma}) - 1},$$

and

$$f(y|\sigma, \mu) = \frac{\sigma}{y} \left(\frac{\log \tau}{\log \mu} \right) \left(\frac{\log y}{\log \mu} \right)^{\sigma - 1} \tau^{(\log y / \log \mu)^\sigma},$$

respectively, where $\mu \in (0, 1)$ is the τ th quantile, $\tau \in (0, 1)$ is assumed known, and $\sigma > 0$ is a shape parameter. For both competitor models, we define systematic components analogous to those in Equations (4) and (5) and set $\tau = 0.5$ to model the median of the response variable.

After evaluating all possible subsets of regressions through the significance of the predictors, AIC, and residual analysis, the systematic components for all classes of regressions are defined by

$$\text{logit}(\mu_t) = \beta_1 + \beta_2\text{GDP}_t + \beta_3\text{HDI}_t + \beta_4\text{DP}_t + \beta_5\text{DGGHE}_t,$$

and

$$\log(\sigma_t) = \gamma_1 + \gamma_2\text{GDP}_t + \gamma_3\text{UP}_t,$$

where $\text{logit}(\mu_t) = \log[\mu_t/(1 - \mu_t)]$ is the logit link function.

The final fitted regressions and their goodness-of-fit measures are reported in Tables 6 and 7, respectively. Table 6 presents the parameter estimates and corresponding *p*-values for the KW, URW, and UW models. In the case of the UW regression, many *p*-values were not statistically significant at the 5% level, capturing only the effects of UP on the response variable MR. The adjusted KW and URW models proved competitive, with all estimates significant at the 5% level, except for the GDP predictor, which was significant at the 10% level. Table 7 shows the results of the Anderson-Darling (AD) test, which we performed to verify the null hypothesis that the quantile residuals are normally distributed. For the UW regression, the AD test rejects the normality hypothesis at the 1% significance level, suggesting that this model is not adequate to the current data. The URW and KW remain as competitive regressions. However, it is noteworthy that our model outperforms the KW for all goodness-of-fit measures. The AIC of the URW regression is the lowest, and the R_G^2 is also in favor of the proposed model, indicating that the URW-fitted model is able to explain about 72.78% of the total variability in the MR. The URW regression model produced the lowest MAE; however, this difference is minimal when considering the scale of MR. To address this limitation and derive a unit-independent metric, the MAE- \bar{MR} ratio was employed, calculated as the quotient of the MAE and the mean MR. The results of this ratio more distinctly demonstrate the advantage of the URW model over the others, underscoring its superiority in terms of predictive accuracy.

Figures 3 and 4 present the diagnostic plots based on the quantile residuals for the fitted URW and KW regressions, respectively. Both fitted regression models, URW and KW, demonstrate suitable results in the graphical analysis of residuals. In the residual plot, the quantile residuals are randomly distributed around zero. In the worm plot, all points lie within the confidence bands and remain close to the central line, with no visible trend. The QQ plot indicates that the sample quantiles are within the limits of the confidence envelopes, suggesting an adequate fit to the data.

Overall, these analyses indicate that the URW regression provides superior fit quality, as evidenced by the lower MAE and MAE- \bar{MR} statistics shown in Table 7, based on the LOOCV approach. These results suggest that the URW regression yields more accurate predictions compared to the KW regression. Therefore, we confirm that the URW model is appropriate and provides better fit quality.

Table 6. Summary of the final fitted regressions for the MR.

Coefficients	KW		URW		UW	
	Estimate	<i>p</i> -Value	Estimate	<i>p</i> -Value	Estimate	<i>p</i> -Value
Intercept (for μ)	−5.9618	0.0421	−6.0836	0.0332	−5.9614	0.6237
GDP	−0.1539	0.0004	−0.1515	0.0006	−0.1520	0.6028
HDI	10.2217	0.0069	10.4728	0.0042	10.4197	0.4536
CHE	0.4216	0.0462	0.4023	0.0468	0.3899	0.6049
DP	−0.3115	0.0000	−0.3082	0.0001	−0.3097	0.1822
DGGHE	−1.0982	0.0004	−1.0861	0.0003	−1.0808	0.3560
Intercept (for σ)	7.0900	0.0000	6.4348	0.0000	6.4599	0.0013
GDP	−0.0431	0.0680	−0.0379	0.0875	−0.0388	0.3108
UP	−0.0718	0.0000	−0.0643	0.0001	−0.0641	0.0014

Table 7. Goodness-of-fit measures for the final fitted regressions.

Regression	AIC	R^2_G	p -Value (AD)	MAE	MAE- \overline{MR}
KW	−69.4907	0.7258	0.7321	0.0556	0.7013
URW	−69.9176	0.7286	0.8473	0.0480	0.6054
UW	−41.4766	−0.1750	0.0002	0.0401	0.5058

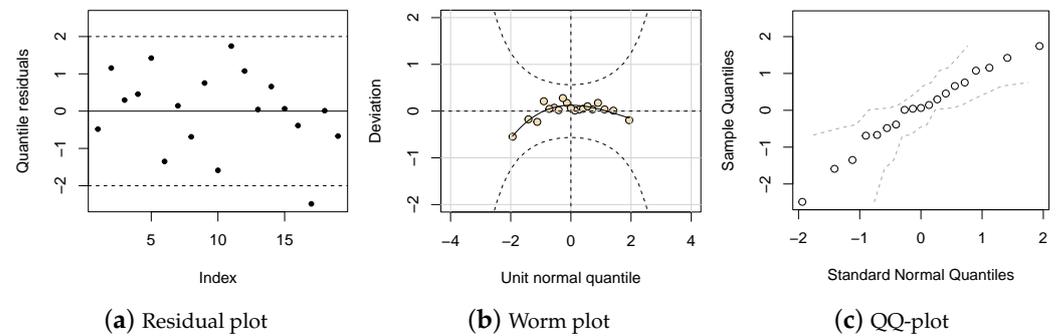


Figure 3. Residuals plots for the fitted URW regression.

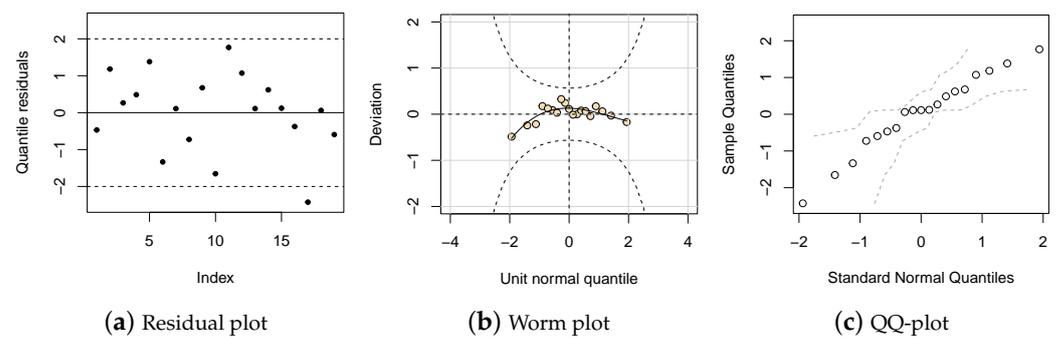


Figure 4. Residuals plots for the fitted KW regression.

The URW model for the median mortality rate due to COVID-19 revealed positive estimated coefficients for HDI (10.4728) and CHE (0.4023) and negative coefficients for GDP (−0.1515), DP (−0.3082), and DGGHE (−1.0861). In terms of variation, a one-unit increase in GDP per capita reduces the median mortality rate by 14.26, while an increase in DGGHE reduces this rate by 66.2%. The DP variable reduces the median rate by 26.5%. On the other hand, an increase in HDI is associated with a significant increase in the median mortality rate, suggesting complex relationships with other factors, while a one-unit increase in the CHE index is associated with a 49.5% increase in the median mortality rate. For the submodel, negative estimated coefficients were found for GDP (−0.0379) and UP (−0.0643). This means that a one-unit increase in GDP per capita and in UP is associated with a reduction of approximately 3.7% and 6.2% in the value of the parameter σ , respectively.

Based on the fitted URW regression, relevant observations were identified regarding the modeling of the median mortality rate due to COVID-19 in Latin American countries. The variables GDP and DGGHE showed negative coefficients, suggesting that increases in GDP and government spending on health infrastructure are associated with a significant reduction in the median mortality rate due to COVID-19. That is, a more robust economy and higher public spending on health seem to enhance resilience against mortality from the disease.

On the other hand, the variable DP showed unexpected results, indicating a negative effect on the response variable. The variables HDI and CHE, also yielding unexpected results, presented positive coefficients, suggesting that increases in the HDI and DGGHE indicators of countries are associated with a higher median mortality rate due to

COVID-19. This counterintuitive interpretation suggests more complex interactions, possibly influenced by other variables such as GDP and DGGHE, which also tend to correlate with economic development, health investment, and diabetes prevalence across countries.

In interpreting the coefficients of the submodel for the parameter σ , both GDP and UP exhibited negative signs, suggesting that increases in GDP and urbanization are associated with a reduction in σ . This indicates that in regions with higher GDP and urbanization, the mortality rate due to COVID-19 tends to accelerate less and stabilize more quickly over time. These results indicate that in areas of higher urban density, the probability of transmission is naturally greater. However, more developed regions with structured economies quickly require the implementation of strict measures, such as social distancing, mask use, and interventions to control the spread of the virus.

5. Concluding Remarks

This article presents a new regression model that explores the relationship between demographic indicators, economic variables, and public health measures with the COVID-19 mortality rate among Latin American countries, a region heavily impacted and regarded as one of the pandemic's epicenters. It is introduced based on the unit ratio-Weibull distribution, which is a helpful tool for modeling random variables in the interval $(0, 1)$, such as rates, proportions, and indices. A general and useful quantile parameterization is introduced to define the new regression model for double-bounded epidemiological data modeling. We defined a systematic structure for the two parameters of the distribution: one represents the quantile of the distribution, and the other, the shape parameter. The parameters were estimated by maximum likelihood, and the performance of the estimators was evaluated through Monte Carlo simulations under different scenarios, considering varying quantile values and finite sample sizes. The URW model was compared with the Kumaraswamy and unit Weibull regressions, proving to be competitive and providing the best fit across various selection criteria and predictive accuracy measures. From the adjusted regression, it was identified that factors such as economic development, Human Development Index, percentage of the urban population, and government investment in health infrastructure are associated with lower COVID-19 mortality rates in Latin American countries. The results indicate that investments in public health and economic infrastructure are essential to reducing the impact of future pandemics and improving public health response policies. The URW regression offers a more robust alternative for capturing the asymmetric and bounded characteristics of mortality rates. This approach provides valuable insights for more effective public policies, helping to understand the impacts of economic and demographic variables on mortality. The ability to apply this methodology to a wide range of fields underscores its versatility, with potential applications in areas such as health, economics, and engineering. Future research could explore its use in analyzing mortality rates from diseases like measles in health, inequality indices like the Gini index and poverty rates in economics, and failure rates or equipment efficiency in engineering. Furthermore, comparing the performance of the proposed model with other approaches across these diverse fields would provide valuable insights into its effectiveness and adaptability.

Author Contributions: Conceptualization, F.A.P.-R. and R.R.G.; methodology, F.A.P.-R., F.J.M.d.A. and R.R.G.; software, F.A.P.-R., F.J.M.d.A. and R.R.G.; validation, F.J.M.d.A., F.A.P.-R. and R.R.G.; formal analysis, F.A.P.-R. and R.R.G.; investigation, F.J.M.d.A., F.A.P.-R. and R.R.G.; resources, F.A.P.-R. and R.R.G.; data curation, F.J.M.d.A.; writing—original draft preparation, F.J.M.d.A.; writing—review and editing, F.A.P.-R. and R.R.G.; visualization, F.J.M.d.A.; supervision, F.A.P.-R. and R.R.G.; project administration, F.A.P.-R. and R.R.G.; funding acquisition, F.A.P.-R. and R.R.G. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by FAPERGS/23/2551-0001595-1, CNPq/306274/2022-1, and Instituto Serrapilheira - 2211-41692, Brazil.

Data Availability Statement: In the article, a set of real data are used, which can be found through the link (<https://github.com/Fernando-code8/URW-Unit-Ratio-Weibull-Regression>- (accessed on 18 October 2024)).

Acknowledgments: We gratefully acknowledge partial financial support from Serrapilheira Institute - Serra - 2211-41692; Fundação de Amparo à pesquisa do Estado do RS - FAPERGS/23/2551-0001595-1. The autor Renata Rojas Guerra acknowledges the Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq/306274/2022-1.

Conflicts of Interest: The funders had no role in the design of the study, the collection, analyses, or interpretation of data, the writing of the manuscript, or in the decision to publish the results.

Abbreviations

The following abbreviations are used in this manuscript:

COVID-19	Coronavirus Disease 2019
URW	Unit Ratio-Weibull
SARS-CoV-2	Severe Acute Respiratory Syndrome Coronavirus 2
GLM	Generalized Linear Models
GAMLSS	Generalized additive Models for Location, Scale, and Shape
MLE	Maximum Likelihood Estimators
AIC	Akaike information criteria
LOOCV	Leave-One-Out Cross-Validation
MAE	Mean Absolute Error
RB%	Relative Bias
MSE	Mean Square Error
CS	Coefficient of Skewness
K	Kurtosis
CR%	Coverage Rate
NRR	Null Rejection Rates
MR	Mortality Rate
KW	Kumaraswamy
UW	unit Weibull
CHE	Current Health Expenditure
DGGHE	Domestic General Government Health Expenditure
DP	Diabetes Prevalence
GDP	Gross Domestic Product per capita
HDI	Human Development Index
UP	Urban Population
CV%	Coefficient of Variation
AD	Anderson-Darling

Appendix A

In the appendix, we provide supplementary results to Section 3. The numerical results of the Monte Carlo simulations for other quantiles of the proposed URW regression are explored below.

Table A1. Results of simulation of URW regression with $\tau = 0.25$.

Measures	n	Scenario 1				Scenario 2			
		$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\gamma}_1$	$\hat{\gamma}_2$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\gamma}_1$	$\hat{\gamma}_2$
RB%	10	0.8955	-12.3577	26.1715	-3.9051	0.9435	-0.9402	17.1284	-3.2589
	15	0.7920	13.0612	19.6379	-2.8108	0.2532	-0.9762	8.6795	17.4380
	30	-0.1949	-2.4211	10.1151	2.8109	0.0268	-0.2949	8.6986	-2.0248
	70	-0.1938	-0.6672	3.1095	4.9114	-0.0034	0.1235	2.9153	3.1665
	150	-0.0813	-0.2715	1.5041	2.1208	-0.0151	-0.0103	1.6613	0.6487
	300	-0.0333	-0.2937	0.6905	1.1025	-0.0036	-0.0082	0.6089	1.0906

Table A1. Cont.

Measures	n	Scenario 1				Scenario 2			
		$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\gamma}_1$	$\hat{\gamma}_2$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\gamma}_1$	$\hat{\gamma}_2$
MSE	10	0.1127	0.5730	3.4659	9.2046	0.0905	0.4523	4.1580	10.6474
	15	0.0980	0.2119	0.6815	2.0895	0.0738	0.1989	0.5999	2.7986
	30	0.0515	0.1199	0.2082	0.6609	0.0371	0.0894	0.3144	0.8253
	70	0.0140	0.0333	0.0522	0.2487	0.0145	0.0271	0.1081	0.3871
	150	0.0069	0.0151	0.0228	0.1185	0.0057	0.0118	0.0399	0.1478
	300	0.0038	0.0085	0.0103	0.0524	0.0027	0.0057	0.0195	0.0765
CS	10	-0.4748	-0.0204	1.1914	0.0696	-0.4858	-0.0093	1.1418	0.1661
	15	-0.4204	0.1240	1.5304	-0.0829	-0.3588	0.0110	1.2062	0.4756
	30	-0.2306	0.0049	1.1835	0.0575	-0.2294	0.0056	0.9943	0.0051
	70	-0.1365	-0.0526	0.6925	0.2627	-0.1321	0.0130	0.6636	0.0908
	150	-0.0336	-0.0321	0.4305	0.2132	-0.1298	-0.0123	0.3829	0.1229
	300	-0.0537	-0.0328	0.2891	0.1429	-0.0664	0.0068	0.3245	0.0706
K	10	3.6347	3.8061	6.3411	4.5570	3.7019	3.8261	6.2545	4.8212
	15	3.4237	3.3723	7.6680	4.4074	3.3342	3.2400	6.0103	4.6433
	30	3.2373	3.3288	5.4315	3.7133	3.2729	3.3480	4.6430	3.4521
	70	3.0450	3.0384	4.0062	3.3648	3.0745	3.1062	3.8198	3.3480
	150	2.9745	2.9595	3.3899	3.1429	3.1177	3.1842	3.3017	3.1480
	300	2.9697	2.9734	3.1233	3.0300	3.0000	3.0280	3.2266	3.1211
CR%	10	0.9000	0.8699	0.9627	0.9622	0.9031	0.8740	0.9597	0.9668
	15	0.9081	0.9019	0.9715	0.9597	0.9229	0.9130	0.9575	0.9632
	30	0.9262	0.9216	0.9552	0.9562	0.9329	0.9272	0.9629	0.9601
	70	0.9363	0.9358	0.9507	0.9505	0.9400	0.9384	0.9558	0.9524
	150	0.9436	0.9446	0.9518	0.9493	0.9499	0.9487	0.9547	0.9525
	300	0.9482	0.9480	0.9505	0.9521	0.9484	0.9499	0.9511	0.9500

Table A2. Simulation results for the quantile residuals of the URW regression with $\tau = 0.25$.

Scenario	n	Mean	Variance	CS	K	NRR10%	NRR5%	NRR1%
1	10	0.0044	1.1585	-0.2039	2.1817	0.0984	0.0390	0.0046
	15	0.0071	1.1059	-0.1692	2.3633	0.1034	0.0451	0.0058
	30	0.0059	1.0488	-0.0759	2.6334	0.0953	0.0435	0.0059
	70	0.0028	1.0210	-0.0354	2.8241	0.0920	0.0442	0.0066
	150	0.0018	1.0092	-0.0158	2.9108	0.0824	0.0407	0.0072
	300	0.0017	1.0043	-0.0098	2.9532	0.0892	0.0427	0.0078
2	10	0.0020	1.1650	-0.2231	2.1788	0.1058	0.0407	0.0048
	15	0.0082	1.1011	-0.1475	2.3926	0.1024	0.0446	0.0061
	30	0.0043	1.0514	-0.0816	2.6208	0.0956	0.0447	0.0065
	70	0.0029	1.0217	-0.0379	2.8211	0.0950	0.0467	0.0060
	150	0.0022	1.0096	-0.0189	2.9110	0.0875	0.0416	0.0081
	300	0.0021	1.0038	-0.0095	2.9534	0.0921	0.0440	0.0087

Table A3. Results of simulation of URW regression with $\tau = 0.75$.

Measures	n	Scenario 1				Scenario 2			
		$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\gamma}_1$	$\hat{\gamma}_2$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\gamma}_1$	$\hat{\gamma}_2$
RB%	10	1.5139	7.0198	59.4137	-15.7750	1.0066	1.0272	50.3627	-15.8088
	15	0.9394	-5.5294	10.4376	58.9365	0.5469	-0.4434	20.8548	19.1723
	30	0.5489	2.1688	12.8607	3.8494	0.3910	0.4465	11.7262	0.5419
	70	0.1713	-0.8342	4.3563	3.9296	0.1220	-0.0153	2.9261	8.5480
	150	0.0951	0.1653	2.2395	1.1441	0.0606	-0.0033	1.8062	2.2326
	300	0.0569	0.2216	0.9944	0.7903	0.0251	-0.0434	0.8836	1.1722

Table A3. Cont.

Measures	n	Scenario 1				Scenario 2			
		$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\gamma}_1$	$\hat{\gamma}_2$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\gamma}_1$	$\hat{\gamma}_2$
MSE	10	0.0762	0.4276	8.8373	19.6653	0.0576	0.3214	12.0601	26.4214
	15	0.0596	0.2170	0.9523	8.5802	0.0470	0.1831	1.9961	8.4663
	30	0.0388	0.1069	0.4495	1.4973	0.0277	0.0780	0.7014	2.0920
	70	0.0097	0.0389	0.1595	0.6060	0.0099	0.0255	0.2353	0.9759
	150	0.0048	0.0139	0.0708	0.2595	0.0040	0.0125	0.0886	0.3329
	300	0.0027	0.0082	0.0250	0.1098	0.0020	0.0062	0.0406	0.1717
CS	10	-0.5182	0.1293	1.3932	-0.1506	-0.4724	0.0761	1.3552	-0.0990
	15	-0.3864	-0.0399	1.1578	0.8260	-0.3172	-0.0624	1.5891	0.0852
	30	-0.1942	0.0519	1.2827	-0.1398	-0.2190	0.0562	1.1600	-0.1521
	70	-0.1104	-0.0524	0.6154	0.0133	-0.1763	0.0369	0.4554	0.1434
	150	-0.1260	0.0217	0.5416	-0.0837	-0.0893	-0.0071	0.4727	0.0027
	300	-0.0580	0.0006	0.2772	0.0145	-0.0400	0.0247	0.3023	0.0043
K	10	3.7395	3.5907	6.7588	4.1142	3.6278	3.5253	6.5626	4.1212
	15	3.5437	3.3600	5.6210	4.9085	3.3445	3.2608	7.9806	4.1577
	30	3.1248	3.1764	6.1000	3.6655	3.2356	3.2399	5.5435	3.6156
	70	3.1267	3.1935	3.6814	3.2058	3.1825	3.1104	3.2776	3.2182
	150	2.9852	2.9519	3.5492	3.1539	3.0646	2.9997	3.3876	3.1305
	300	2.9573	2.9404	3.0739	3.0106	2.9913	3.0221	3.1717	3.0464
CR%	10	0.8427	0.8475	0.9389	0.9402	0.8391	0.8450	0.9386	0.9413
	15	0.8503	0.8496	0.9033	0.9073	0.8815	0.8780	0.9191	0.9157
	30	0.9100	0.9174	0.9410	0.9389	0.9104	0.9190	0.9408	0.9375
	70	0.9372	0.9387	0.9429	0.9465	0.9357	0.9361	0.9443	0.9458
	150	0.9393	0.9389	0.9495	0.9474	0.9417	0.9393	0.9462	0.9450
	300	0.9494	0.9498	0.9523	0.9526	0.9479	0.9495	0.9504	0.9489

Table A4. Simulation results for the quantile residuals of the URW regression with $\tau = 0.75$.

Scenario	n	Mean	Variance	CS	K	NRR10%	NRR5%	NRR1%
1	10	-0.0089	1.1467	-0.1145	2.2479	0.0909	0.0422	0.0058
	15	-0.0023	1.0776	0.0049	2.4589	0.0862	0.0372	0.0037
	30	-0.0034	1.0386	0.0069	2.6568	0.0846	0.0385	0.0053
	70	-0.0007	1.0166	0.0052	2.8431	0.0878	0.0398	0.0066
	150	-0.0005	1.0080	0.0016	2.9171	0.0863	0.0400	0.0061
	300	-0.0003	1.0042	-0.0001	2.9565	0.0921	0.0450	0.0078
2	10	-0.0078	1.1472	-0.1131	2.2581	0.0886	0.0411	0.0061
	15	-0.0055	1.0830	-0.0146	2.4164	0.0852	0.0372	0.0043
	30	-0.0035	1.0399	0.0018	2.6503	0.0854	0.0374	0.0050
	70	-0.0009	1.0173	0.0022	2.8406	0.0879	0.0388	0.0075
	150	-0.0004	1.0079	0.0026	2.9179	0.0873	0.0402	0.0071
	300	-0.0002	1.0043	-0.0005	2.9567	0.0899	0.0402	0.0071

References

- World Health Organization (WHO). Indicators. 2024. Available online: <https://www.who.int/emergencies/diseases/novel-coronavirus-2019> (accessed on 13 October 2024).
- Ashmore, P.; Sherwood, E. An overview of COVID-19 global epidemiology and discussion of potential drivers of variable global pandemic impacts. *J. Antimicrob. Chemother.* **2023**, *78*, ii2–ii11. [CrossRef]
- Ashktorab, H.; Pizuorno, A.; Oskroch, G.; Fierro, N.A.; Sherif, Z.A.; Brim, H. COVID-19 in Latin America: Symptoms, Morbidities, and Gastrointestinal Manifestations. *Gastroenterology* **2021**, *160*, 938–940. [CrossRef] [PubMed]
- Li, J.; Lai, S.; Gao, G.F.; Shi, W. The emergence, genomic diversity and global spread of SARS-CoV-2. *Nature* **2021**, *600*, 408–418. [CrossRef] [PubMed]
- Cimerman, S.; Chebabo, A.; Cunha, C.A.D.; Rodríguez-Morales, A.J. Deep impact of COVID-19 in the healthcare of Latin America: The case of Brazil. *Braz. J. Infect. Dis.* **2020**, *24*, 93–95. [CrossRef] [PubMed]
- Anderson, R.M.; May, R.M. *Infectious Diseases of Humans: Dynamics and Control*; Oxford University Press: Oxford, UK, 1991.

7. Kermack, W.O.; McKendrick, A.G. A contribution to the mathematical theory of epidemics. *Proc. R. Soc. Lond. Ser. A Contain. Pap. Math. Phys. Character* **1927**, *115*, 700–721.
8. Fauzi, N.F.; Bakhtiar, N.S.A.; Khairudin, N.I.; Halim, H.Z.A.; Shafii, N.H. A Performance of SIR Model in Predicting the Number of COVID-19 Cases in Malaysia based on Different Phase of COVID-19 Outbreak. *J. Comput. Res. Innov.* **2023**, *8*, 75–83. [[CrossRef](#)]
9. Kodera, S.; Rashed, E.A.; Hirata, A. Correlation between COVID-19 morbidity and mortality rates in Japan and local population density, temperature, and absolute humidity. *Int. J. Environ. Res. Public Health* **2020**, *17*, 5477. [[CrossRef](#)] [[PubMed](#)]
10. Ozkan, A.; Ozkan, G.; Yalaman, A.; Yildiz, Y. Climate risk, culture and the Covid-19 mortality: A cross-country analysis. *World Dev.* **2021**, *141*, 105412. [[CrossRef](#)] [[PubMed](#)]
11. Leffler, C.T.; Ing, E.; Lykins, J.D.; Hogan, M.C.; McKeown, C.A.; Grzybowski, A. Association of country-wide coronavirus mortality with demographics, testing, lockdowns, and public wearing of masks. *Am. J. Trop. Med. Hyg.* **2020**, *103*, 2400–2411. [[CrossRef](#)] [[PubMed](#)]
12. Liang, L.L.; Tseng, C.H.; Ho, H.J.; Wu, C.Y. COVID-19 mortality is negatively associated with test number and government effectiveness. *Sci. Rep.* **2020**, *10*, 12567. [[CrossRef](#)] [[PubMed](#)]
13. McCullagh, P.; Nelder, J.A. *Generalized Linear Models*, 2nd ed.; Chapman and Hall: London, UK, 1989.
14. Rigby, R.A.; Stasinopoulos, D.M. Generalized additive models for location, scale and shape. *J. R. Stat. Soc. Ser. C Appl. Stat.* **2005**, *54*, 507–554. [[CrossRef](#)]
15. Guerra, R.R.; Peña-Ramírez, F.A.; Ribeiro, T.F.; Cordeiro, G.M.; Mafalda, C.P. Unit Regression Models to Explain Vote Proportions in the Brazilian Presidential Elections in 2018. *Rev. Colomb. Estadística* **2024**, *47*, 283–300.
16. Lemonte, A.J.; Bazán, J.L. New class of Johnson distributions and its associated regression model for rates and proportions. *Biom. J.* **2016**, *58*, 727–746. [[CrossRef](#)] [[PubMed](#)]
17. Mazucheli, J.; Menezes, A.F.B.; Fernandes, L.B.; de Oliveira, R.P.; Ghitany, M.E. The unit-Weibull distribution as an alternative to the Kumaraswamy distribution for the modeling of quantiles conditional on covariates. *J. Appl. Stat.* **2020**, *47*, 954–974. [[CrossRef](#)] [[PubMed](#)]
18. Peña-Ramírez, F.A.; Guerra, R.R.; Mafalda, C.P. The unit ratio-extended Weibull family and the dropout rate in Brazilian undergraduate courses. *PLoS ONE* **2023**, *18*, e0290885. [[CrossRef](#)] [[PubMed](#)]
19. Rigby, R.; Stasinopoulos, D. The GAMLSS project: A flexible approach to statistical modelling. In *New Trends in Statistical Modelling, Proceedings of the 16th International Workshop on Statistical Modelling, Odense, Denmark, 2–6 July 2001*; Statistical Modelling Society: Amsterdam, The Netherlands, 2001; Volume 337.
20. Stasinopoulos, D.M.; Rigby, R.A. Generalized additive models for location scale and shape (GAMLSS) in R. *J. Stat. Softw.* **2007**, *23*, 1–46. [[CrossRef](#)]
21. Stasinopoulos, M.D.; Rigby, R.A.; Heller, G.Z.; Voudouris, V.; De Bastiani, F. *Flexible Regression and Smoothing: Using GAMLSS in R*; CRC Press: Boca Raton, FL, USA, 2017.
22. Dunn, P.K.; Smyth, G.K. Randomized quantile residuals. *J. Comput. Graph. Stat.* **1996**, *5*, 236–244. [[CrossRef](#)]
23. Pereira, G.H. On quantile residuals in beta regression. *Commun. Stat.-Simul. Comput.* **2019**, *48*, 302–316. [[CrossRef](#)]
24. Queiroz, F.F.; Lemonte, A.J. A broad class of zero-or-one inflated regression models for rates and proportions. *Can. J. Stat.* **2020**, *49*, 566–590. [[CrossRef](#)]
25. Buuren, S.V.; Fredriks, M. Worm plot: A simple diagnostic device for modelling growth reference curves. *Stat. Med.* **2001**, *20*, 1259–1277. [[CrossRef](#)] [[PubMed](#)]
26. Nagelkerke, N.J. A note on a general definition of the coefficient of determination. *Biometrika* **1991**, *78*, 691–692. [[CrossRef](#)]
27. Akaike, H. A new look at the statistical model identification. *IEEE Trans. Autom. Control* **1974**, *19*, 716–723. [[CrossRef](#)]
28. James, G.; Witten, D.; Hastie, T.; Tibshirani, R.; Taylor, J. *An Introduction to Statistical Learning*; Springer: Berlin/Heidelberg, Germany, 2013; Volume 112.
29. Bayes, C.L.; Bazán, J.L.; De Castro, M. A quantile parametric mixed regression model for bounded response variables. *Stat. Its Interface* **2017**, *10*, 483–493. [[CrossRef](#)]
30. The World Bank. Indicators. 2021. Available online: <https://data.worldbank.org/indicator> (accessed on 19 December 2020).
31. Adeleye, B.N.; Gershon, O.; Ogunlape, A.; Owolabi, O.; Ogunrinola, I.; Adediran, O. Comparative investigation of the growth-poverty-inequality trilemma in Sub-Saharan Africa and Latin American and Caribbean Countries. *Heliyon* **2020**, *6*, e05631. [[CrossRef](#)]
32. Conceição, P. *Human Development Report 2020—The Next Frontier: Human Development and the Anthropocene*; Human Development Report Series; United Nations Development Programme: New York, NY, USA, 2020.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.